Haffez H, Chisholm DR, Valentine R, Pohl E, Redfern C, Whiting A.

The molecular basis of the interactions between synthetic retinoic acid analogues and the retinoic acid receptors.

***MedChemComm* 2017**

https://doi.org/10.1039/C6MD00680A

MedChemComm

ARTICLE

The molecular basis of the interactions between synthetic retinoic acid analogues and the retinoic acid receptors

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org

Hesham Haffez,[a,b,c,d] David R. Chisholm,[a] Roy Valentine,[e] Ehmke Pohl,[b] Christopher Redfern[c] and Andrew Whiting[a]

All-trans-retinoic acid (ATRA) and its synthetic analogues EC23 and EC19 direct cellular differentiation by interacting as ligands for the retinoic acid receptor (RARα, β and γ) family of nuclear receptor proteins. To date, a number of crystal structures of natural and synthetic ligands complexed to their target proteins have been solved, providing molecular level snap-shots of ligand binding. However, a deeper understanding of receptor and ligand flexibility and conformational freedom is required to develop stable and effective ATRA analogues for clinical use. Therefore, we have used molecular modelling techniques to define RAR interactions with ATRA and two synthetic analogues, EC19 and EC23, and compared their predicted biochemical activities to experimental measurements of relative ligand affinity and recruitment of coactivator proteins. A comprehensive molecular docking approach that explored the conformational space of the ligands indicated that ATRA is able to bind the three RAR proteins in a number of conformations with one extended structure being favoured. In contrast the biologically-distinct isomer, 9-*cis*-retinoic acid (9CRA), showed significantly less conformational flexibility in the RAR binding pockets. These findings were used to inform docking studies of the synthetic retinoids EC23 and EC19, and their respective methyl esters. EC23 was found to be an excellent mimic for ATRA, and occupied similar binding modes to ATRA in all three target RAR proteins. In comparison, EC19 exhibited an alternative binding mode which reduces the strength of key polar interactions in RARα/γ but is well-suited to the larger RARβ binding pocket. In contrast, docking of the corresponding esters revealed the loss of key polar interactions which may explain the much reduced biological activity. Our computational results were complemented using an *in vitro* binding assay based on FRET measurements, which showed that EC23 was a strongly binding, pan-agonist of the RARs, while EC19 exhibited specificity for RARβ, as predicted by the docking studies. These findings can account for the distinct behaviour of EC23 and EC19 in cellular differentiation assays, and additionally, the methods described herein can be further applied to the understanding of the molecular basis for the selectivity of different retinoids to RARα, β and γ.

Introduction

All-*trans*-retinoic acid (ATRA), the major metabolite of vitamin A, has essential roles in many biological processes during embryogenesis and homeostasis in chordates, including cell differentiation, proliferation and apoptosis, embryonic development and vision.[1,2] Due to this powerful and wide ranging biological activity, ATRA and its isomers have been used as therapeutics for the treatment of various cancers and skin conditions.[3,4] In addition, ATRA is routinely used *in vitro* as an inducer of cellular differentiation; however, its usage has limitations, as the compound readily isomerises to mixtures of ATRA, 9-*cis*-retinoic acid (9CRA), 13-*cis*-retinoic acid (13CRA) and other species.[5,6] This instability is derived from the five conjugated double bonds that absorb visible light at wavelengths of 300-400 nm. In addition, ATRA is inactivated *in vivo* by inducible and constitutively expressed cytochrome P450 enzymes.[7,8] While 9CRA may also have a cellular role as a ligand to activate transcription factors, a natural physiological function for 13CRA as a transcription-factor ligand is less certain.[5]

We have previously described two light-stable synthetic retinoid analogues, EC23 and EC19 (Figure 1, left),[9] differing from each other only in the position of the carboxylic acid group on the terminal substituted ring. EC23 was designed to mimic ATRA, whereas the *meta*-position carboxylic acid of EC19 changes the effective geometry of the molecule in a similar manner to the alkene geometry differences between ATRA and 13CRA. Previous studies showed that EC23 induced neuronal differentiation in pluripotent stem cells at a lower

[a] Department of Chemistry Durham University, South Road, Durham, DH1 3LE, UK.
[b] Department of Biosciences & Biophysical Sciences Institute Durham University, South Road, Durham DH1 3LE, UK.
[c] Northern Institute for Cancer Research, Medical School, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK.
[d] Department of Biochemistry and Molecular Biology, Pharmacy College, Helwan University, Cairo, Egypt.
[e] High Force Research Ltd., Bowburn North Industrial Estate, Bowburn, Durham, DH6 5PF, UK.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

MedChemComm

ARTICLE

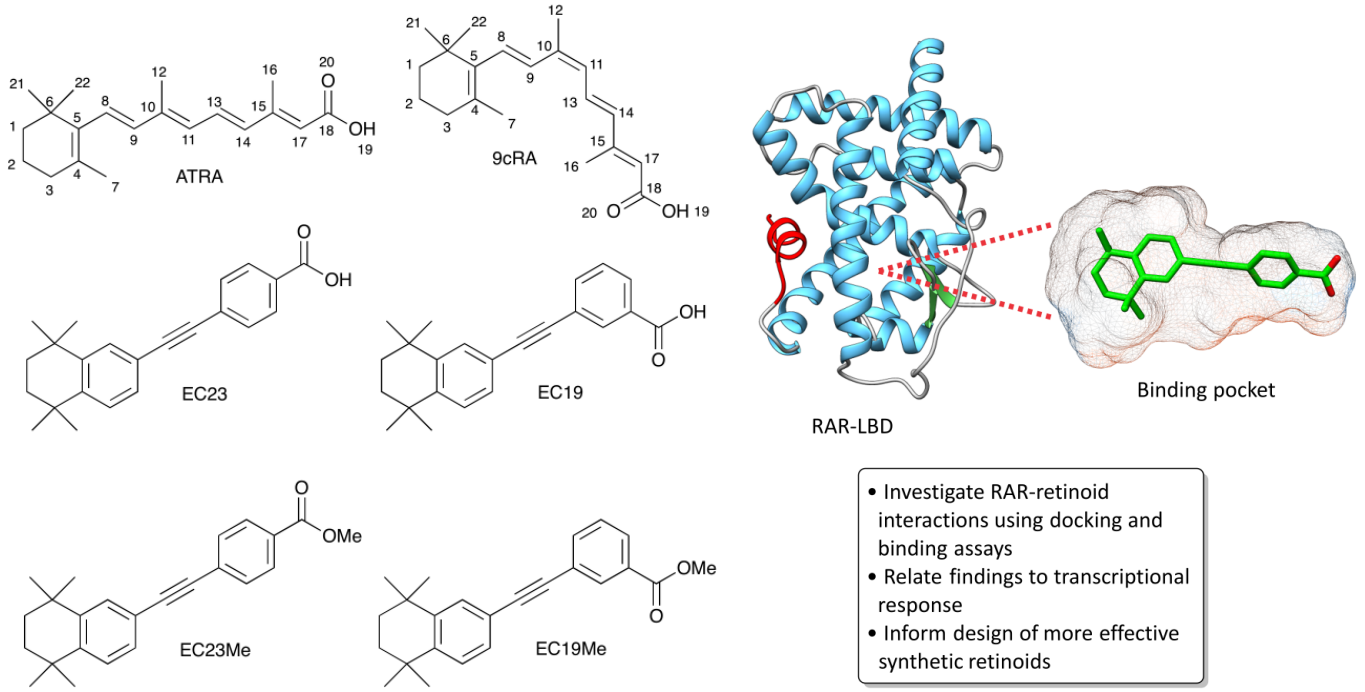

Figure 1: Left, chemical structures of endogenous retinoids ATRA and 9CRA, synthetic retinoids EC23 and EC19, and synthetic retinoid esters EC23Me and EC19Me.[3,9] Right, proposed investigation into the interactions of natural and synthetic retinoids by combining *in silico* docking studies with an *in vitro* binding assay.

concentration than ATRA, whereas, EC19 facilitated epithelial-like development, rather than neuronal differentiation.[9] Thus, although EC23 and EC19 are chemically similar, they have substantially different biological effects.

The biological activity of ATRA is the result of its interactions with the retinoic acid receptors (RARs), which are ligand-dependent, transcriptional regulators. Three closely related isotypes, designated RARα, RARβ and RARγ, have been identified in vertebrates.[2,10,11] The interaction of ATRA with the ligand-binding domain (LBD) of these receptors results in conformational changes, particularly with respect to the C-terminal helix 12 (H12, shown in red in Figure 1), which facilitates recruitment of transcriptional co-activators, leading to DNA binding and hence gene activation.[11,12] Essentially, the retinoid enters the ligand binding pocket of the RAR LBD, guided by an electrostatic gradient generated by a cluster of polar residues including a positively charged arginine at the bottom of the pocket.[13] The carboxylate of the retinoid then associates with this polar cluster, anchoring the retinoid inside the pocket, whereupon H12 is thought to swing around from its solvent exposed *apo*-position to close the pocket, trapping the retinoid inside.[13,14] Previous X-ray crystal structures of RAR-retinoid complexes have shown that the retinoids are completely buried within the binding pocket formed by residues from helices H3, H5, H7 and H11, and a β-turn, in a mainly α-helical binding domain. The majority of contacts within the LBD are van der Waals interactions, except for the carboxylate group of the retinoid which forms a salt bridge with a conserved Arg residue from helix H5 and a network of hydrogen bonds with a Ser residue (β-turn) and/or Lys residues.[13,14] The retinoid-bound conformation provides a stable platform for the recruitment and binding of coactivator peptides and dimerization to the related retinoid-X-receptors (RXR).[15–17]

Sequence alignment of RARα, β and γ shows that only three residues in the binding pocket are variable, *i.e.* RARγ exhibits Ala234 (Ser232 in α and Ala225 in β), Met272 (Ile270 in α and Ile263 in β) and Ala397 (Val395 in α and Val388 in β).[18,19] These divergent residues account for the observed selectivity of the different retinoic acids towards binding with RARα, β and γ, and have been exploited to inform the design of a variety of isotype-specific synthetic retinoids.[3,20,21]

The second endogenous ligand for the RARs is 9CRA, a *cis*-isomer of ATRA, which acts as an agonist for both RARs and a second class of retinoic acid receptor, the retinoid X-receptors, (RXR).[22–24] Crystallographic studies of the LBDs of both types of receptor show that although the overall architecture of RARs and RXRs is very similar, the shape of the ligand binding pocket

differs markedly between RARs and RXRs.[25] RARs possess a linear "I" shaped pocket, whereas the RXRs exhibit a shorter and more restrictive "L" shape. As a result, the linear retinoid ATRA mainly acts as a ligand for RARs while 9CRA exhibits high binding affinity for both the RARs and RXRs.[22,26,27]

Although EC23 is more effective than either ATRA or EC19 in differentiation assays,[9] cell-based biological assays cannot be used to elucidate the affinities of synthetic retinoids for RAR receptor types because the specificities of different RARs as drivers of biological responses are poorly understood. Therefore, the molecular basis of the ligand-receptor interactions that underpin these differences in biological activity remains unclear. Transcriptional responses mediated by retinoids involve a complex series of events, starting with ligand binding of the nuclear receptor and leading to the recruitment of co-activator proteins. It is, therefore, important to understand the factors that affect retinoid binding to RARs, especially as a function of the chemical structure and conformational flexibility of the ligand in the context of each RAR type. To address this question, receptor-binding assays are required to confirm predictions from structural modelling studies and to explain how structural differences constrain the biological activity of different ligands. Accordingly, the aim of this study (Figure 1, right) was to unravel the molecular interactions of these retinoids with the RARs by modelling the binding of EC23 and EC19, and their respective methyl esters, in comparison to ATRA and 9CRA to the ligand-binding pocket of each RAR isotype, and to compare the results with experimental measurements of ligand-receptor binding interactions. This approach will enhance our understanding of the molecular basis of retinoid-mediated cellular activity and help inform the design of the next generation of synthetic retinoids as reagents, therapeutics and biological probes.

Experimental

Molecular Modelling

Molecular structures of all compounds (ATRA, 9CRA, EC23, EC19, EC23Me and EC19Me) were generated using Spartan '14 (Wavefunction Inc., Irvine, CA).[28] These were minimised using a molecular mechanics force field, followed by semi-empirical molecular orbital (AM1) methods to generate a conformer distribution *in vacuo*. The generated conformations of each compound were then re-minimised using a higher level of theory (Hartree-Fock, 3-21G) to ensure that the generated conformer distribution was realistic. All possible conformations of all compounds were saved as .mol2 files for use in the docking studies.

Docking

Due to the significant conformational flexibility exhibited by ATRA in particular, and the close energy differences between the observed conformers, we were interested to examine how a docking simulation would interpret the binding process from a variety of starting points.[29] Accordingly, rather than docking only the lowest energy structures, all of the output conformations from the conformer distribution calculations of ATRA, 9CRA, EC23, EC23Me, EC19 and EC19Me were prepared for docking using GOLD.[30]

Crystal structures were retrieved from the RCSB protein data bank as PDB files (3KMR[31] for RARα, 1XAP[20] for RARβ and 2LBD[13] for RARγ). The bound ligands were removed from the structures, and used only as a positional reference. Hydrogen atoms were added to protein residues using the default GOLD settings[30]. All solvent molecules were removed. Active site residues were selected within a diameter of 15 Å, measured from a selected point at the centre of the ligand position, and therefore included the entire binding pocket. It is important to note that the carboxylic acid moiety was considered as a carboxylate, rather than protonated, as this is more realistic with respect to physiological pH, and to the likelihood that the negative charge is stabilised by the closely positioned polar residues in the binding pocket.[15,32] ChemScore was chosen as the most appropriate target function in the genetic algorithm to balance between computational speed and the reliability of the GOLD predictions for the possible conformations of the different retinoids binding to RARs.[33] The genetic algorithm parameters were based on previous examples of docking hydrophobic ligands: population size 100; number of islands 5; niche size 2; selection pressure 1.1; migrate 2; and number of operators 100,000.[34] A search efficiency of 200% was used, which dictates maximum ligand flexibility. Ligands were also allowed maximum flexibility. During the docking process no limit was placed on the number of binding poses retained, though typically 3-10 solutions were retained by the genetic algorithm.

In vitro binding assay

The binding interactions of ATRA, 9CRA, EC23, EC23Me, EC19 and EC19Me with RARα, β and γ were determined *in vitro* by time-resolved fluorescence resonance energy transfer (TR-FRET)[35] using the LanthaScreen TR-FRET RARα, RARβ, and RARγ co-activator assays according to the manufacturer's (Invitrogen™) instructions. This uses a terbium-labelled anti-GST antibody, a fluorescein-labelled coactivator peptide, and RARα, β or γ ligand-binding domains fused to glutathione-S-transferase (GST) in a homogenous assay format. All experiments were performed in black, 384-well low-volume plates in the dark at room temperature with a 4 h incubation time. The final assay volume was 20 µL and all dilutions were carried out using TR-FRET assay buffer, with a final DMSO concentration of 1%. A mixture of either 3.5 nM GST-RARα-LBD or 2.5 nM GST-RARβ-LBD or 3 nM GST-RARγ-LBD with 62.5 nM Tb-anti GST antibody, 30 µM fluorescein-labelled peptide (Fluorescein-D22, Fluorescein-SRC2-2 and Fluorescein-PGC1a for RAR-α, RAR-β, and RAR-γ respectively), and retinoid or DMSO control was added to each of the wells. Each ligand assay was performed in duplicate and measured using a PHERAstar FS Microplate Reader (BMG Labtech, Ortenberg, Germany) with instrument settings as described in the manufacturer's instructions for LanthaScreen assays. The TR-FRET signal was expressed as the ratio of the signals at 520 nm

and 490 nm. The data were fitted to a three-parameter ligand-binding curve using SigmaPlot (version 12.5, Systat Software Inc., San Jose CA) and normalized to the lower asymptote of each binding curve. The TR-FRET binding assay produces sigmoidal curves with different mid-point (EC50) and upper asymptote. Since the assay involves interactions between ligand and LBD, and ligand-dependent interactions between the LBD and fluorescein-labelled coactivator, the binding assay data were interpreted by reference to simulations of these coupled chemical reactions using the biochemical simulator COPASI version 4.3.[36]

Results and Discussion

Molecular modelling and docking

From the conformation studies, it was found that ATRA and 9CRA exhibited a range of accessible low energy conformations, all within around 3-5 kcal/mol, while the synthetic retinoids EC23, EC19, EC23Me and EC19Me exhibited between 1-4 conformations (see ESI for further detail on all calculated conformations). For ATRA, the presence of the methyl groups on the cyclohexenyl ring resulted in low energy conformations in which the conjugated polyene was twisted out of plane with the 6-membered ring. An all s-*trans*-polyene configuration for ATRA was found as the lowest energy conformation, which was within 1 kJ/mol of the second lowest general structure, in which ATRA adopted one s-*cis* sigma-bond rotation (C_8=C_9) in the middle of the polyene, apparently relieving strain due to the presence of the C_7 methyl group. All low energy conformations for ATRA showed an out-of-plane twisting of the polyene from the plane of the cyclohexenyl ring, especially the endo-cyclic alkene.[29,37] Otherwise, the conformations only differed in terms of which sigma bonds exhibited s-*cis* conformations (see SI for structures). The difference in energy between the lowest and highest energy conformations was only around 15 kJ/mol and hence, all are selectable as possible structures that could bind to the receptor sites since all such conformations would exist in similar amounts in solution at RT and physiological temperature.

For 9CRA, the lowest energy conformation showed similar trends to that of ATRA, with the polyene section twisting out of plane from the cyclohexenyl ring due to steric repulsion with the 6-ring *geminal* dimethyl groups (C_7 in particular); similar *s-cis* sigma bond rotational effects were also present. The difference between the lowest and highest energy conformations was again only around 19 kJ/mol. EC23, EC19 and their respective methyl esters displayed significantly less conformational flexibility as expected due to their more restrained chemical structures and particularly the linear alkyne motif. In EC23, for example, the plane of the benzoic acid polar region was perpendicular with respect to the tetrahydronaphthalene hydrophobic region. In EC23Me two low energy structures were obtained with each exhibiting opposing ring conformations for the tetrahydronaphthalene ring. Only one solution was returned for EC23 and EC19, suggesting that the energy difference between the relative orientations of the two rings was negligible.

After the possible starting conformations of each retinoid were docked into the RARs, each solution was examined before further analysis, to ensure that the archetypal positioning of the retinoid ligand according to the existing literature crystal structures was correctly occupied. The vast majority of solutions satisfied this criterion, though in the case of 9CRA some solutions indicated atypical poses, and, accordingly, exhibited significantly lower fitness scores; these solutions were discarded from further analysis. Solutions that exhibited the expected position were close in fitness score (+/- 1.5 fitness score). Each binding pose was then extracted from the protein structure, and individually examined to assign and assess its conformation. Importantly, these binding poses were independent of the starting conformation (by inspection), thus indicating that the docking protocol sampled the conformational space sufficiently and predicts binding poses that were determined by the structure of the RAR binding pockets. Due to their low conformational flexibility, the docking solutions of EC23, EC23Me, EC19 and EC19Me exhibited only minor differences between different binding poses, principally in the angle between the plane of the tetrahydronaphthalene hydrophobic region with respect to the plane of the benzoate/benzoic acid polar region. 9CRA exhibited only two conformations (Figure 2) with the expected 9-*cis* polyene structure with two possible orientations of the cyclohexenyl ring.[37] In contrast, analysis of the docking solutions of ATRA identified eleven possible binding conformations. To gain further insight into the distribution of these conformations and to ascertain whether a particular conformation was favoured, a clustering approach was developed whereby each docking pose of ATRA and 9CRA was individually examined and assigned to a particular conformation, and the percentage incidence of each was plotted for each RAR as a histogram in Figures 2-3. Conformations 1, 2 and 7 were the most frequently obtained by GOLD. Furthermore, conformation 1 was found at around double the frequency of conformations 2 and 7. Interestingly, conformation 1 was only the second lowest energy conformation according to the initial conformer distribution calculations (conformation 7 was the lowest energy), indicating that GOLD correctly attempted to fit the most suited conformation to the shape of the binding pocket and did not simply bind the lowest energy conformation. The less sampled conformations were also those conformations with higher energy in the initial conformational distribution calculations. These observations further highlighted the fine differences in energy between each of the possible conformations of ATRA, and initially indicated that ATRA likely binds in a variety of conformations, although one may be favoured in particular.[38] To confirm that the less-frequently sampled conformations were not simply derived from solutions that had not completely converged, examples of each docked conformation was extracted from the protein structure and docked again, further analysis of the distribution of binding poses in Figure 3

indicated that the range of obtained conformations agreed with the known shapes and structures of the respective RARα, β and γ binding sites. For example, docking ATRA into the larger pocket of RARβ resulted in a much greater range of conformations than the slightly smaller, thinner pockets of RARα and RARγ.[20,39] This suggests that the RARβ pocket was also able to select the less linear conformations, such as conformations 4, 6 and 10. RARα and RARγ mainly accommodated the more linear conformations, such as conformation 1, 2 and 7, during the docking process.

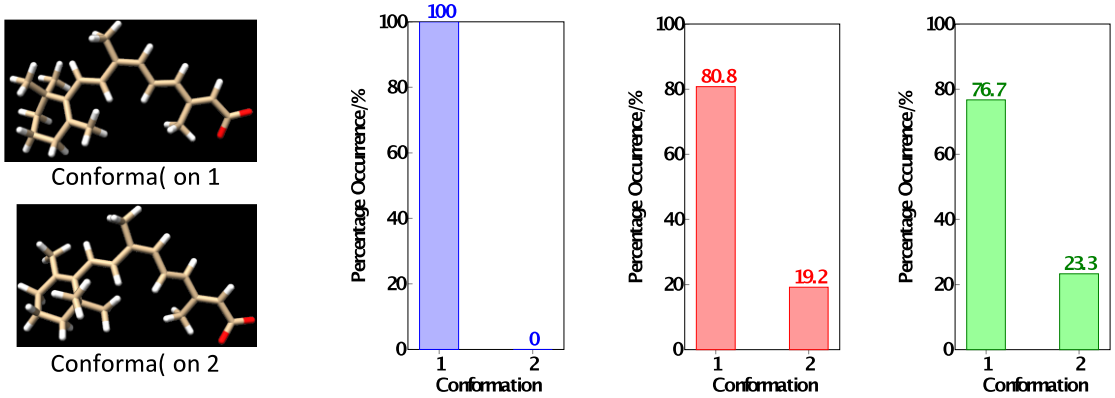

Figure 2: (Left) The conformations of 9CRA observed during docking into RARα, RARB and RARγ, in order of their observation. (Right) Conformation clustering histograms summarising the distribution of the conformations of 9CRA that were observed in docking to RARα (blue), RARβ (red) and RARγ (green). 53 potential docking solutions were examined in RARα (11 discarded due to atypical positioning), 78 potential docking solutions were examined in RARβ (five discarded due to atypical positioning), and 73 potential docking solutions were examined in RARγ (none discarded).

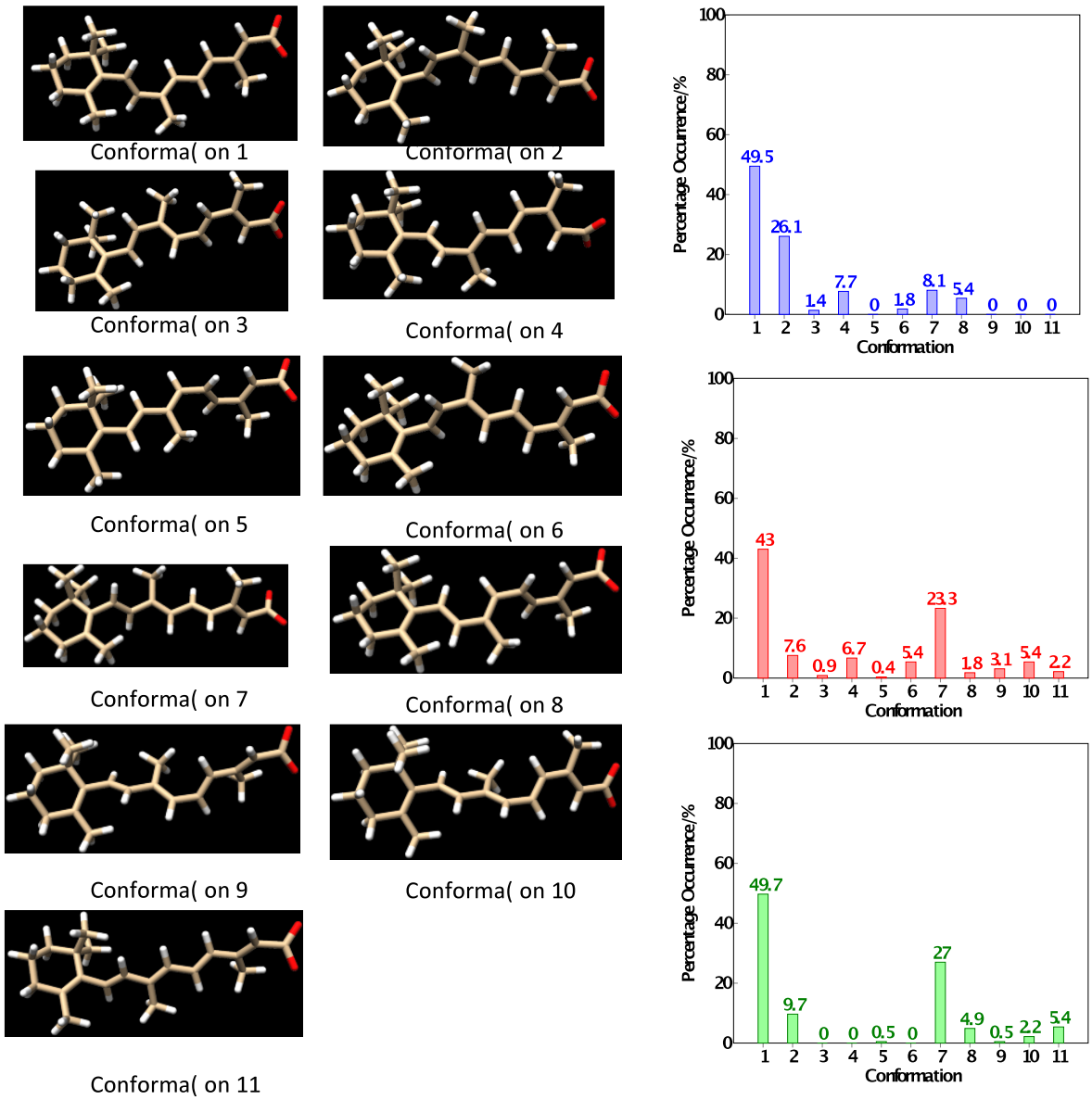

Figure 3: (Left) The conformations of ATRA observed during docking into RARα, RARB and RARγ with respect to the alkene stereochemistry, in order of their observation. (Right) Conformation clustering histograms summarising the distribution of the conformations of ATRA that were observed in docking to RARα (blue), RARβ (red) and RARγ (green). 222 potential docking solutions were examined in RARα, 223 potential docking solutions were examined in RARβ, and 185 potential docking solutions were examined in RARγ.None of these solutions were discarded due to unrealistic positioning. The conformation frequency distribution profiles differ significantly between each receptor (Fisher's exact test, P<0.001).

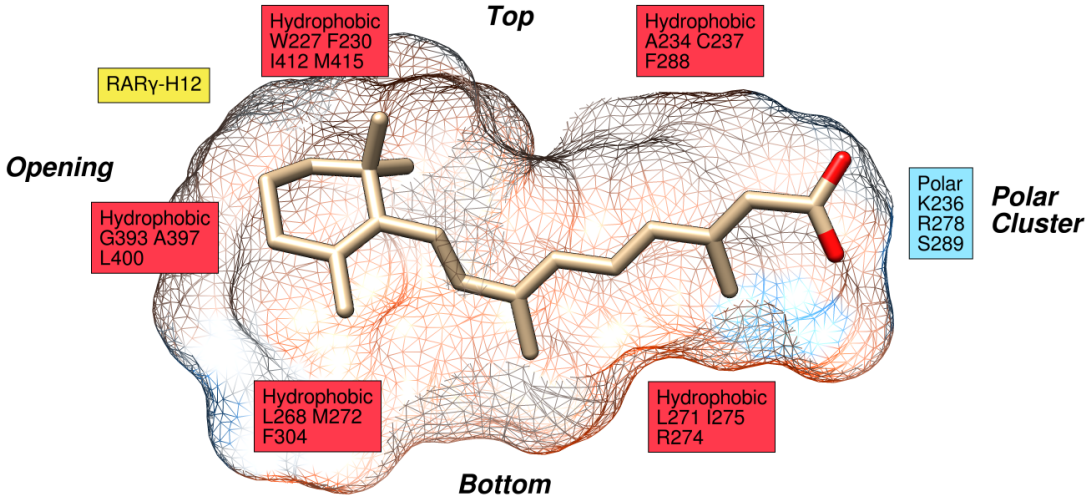

Figure 4 Ligand interaction diagram of the most frequent encountered ATRA conformation, ATRA-Conformation 1, in the RARγ (PDB: 2LBD) binding pocket (mesh),[13] highlighting the predominantly hydrophobic nature of the binding of retinoids to RARα, -β and -γ. A cluster of polar residues (K236, R278 and S289 in the case of RARγ) at the base of the pocket anchor the retinoid via the carboxylate. Helix 12 (H12, in yellow) is positioned in the holo conformation, covering the opening to the pocket (hydrogen -atoms were omitted for clarity in this and all other figures).

Figure 4 shows the structural features of the RAR binding pocket (ATRA conformation 1 in RARγ is shown as an example), and highlights the mildly twisted ligand conformation, the predominantly hydrophobic nature of the pocket, and the small polar cluster at the end that associates with the retinoid carboxylate. The *gem*-dimethyl moiety of the 6-ring (C_{21}/C_{22}) is closely associated with the rich hydrophobicity at the top of the pocket, and the two alkenyl methyl groups (C_{12} & C_{16}) are well positioned for hydrophobic contacts with the bottom of the pocket. The opening of the pocket is enclosed by H12 which traps the retinoid inside and interacts with the hydrophobic region of the retinoid.[13,14,31] A strong hydrophobic interaction is required here to provide a stable platform for the recruitment of coactivator peptides to the outer surface of H12. The carboxylate is shown near to the polar cluster allowing strong polar interactions with the conserved Arg and Ser and Lys236 in this region. The slight twist in overall ligand shape first observed in the conformational analysis agrees with the shape of the RARγ binding pocket, and accommodates the sterically demanding *gem*-dimethyl group into the widest part of the pocket.

The observed conformational distribution of ATRA in RARγ was particularly interesting when comparing the docking solutions of ATRA with that of the experimental crystal structure in the literature.[13,29] Figure 5 shows a superposition of the highest scoring poses of the most sampled conformations in RARγ (conformations 1 & 7) with the crystal structure (PDB: 2LBD), highlighting a generally good agreement. The difference in 6-ring positioning between conformation 1 & 7 is due to the differences in alkene conformation between the C_8-C_9 double bond adjacent to the ring, and the polyene chain. When these contrasting positions are compared with the 2LBD ligand position, the position in the crystal structure can essentially be described as an average of the two binding poses. This is particularly relevant, since given the resolution of the crystal structure (2.06 Å) and the high conformational flexibility of the ligand, it would have undoubtedly been difficult to unequivocally resolve multiple conformations of ATRA. This example therefore highlights the potential benefits of using a docking analysis such as that described here to understand the accessible ligand conformations in a binding site, and to use the resulting information to better understand the fitting of a ligand to electron density from X-ray diffraction, particularly when dealing with medium resolution diffraction data.

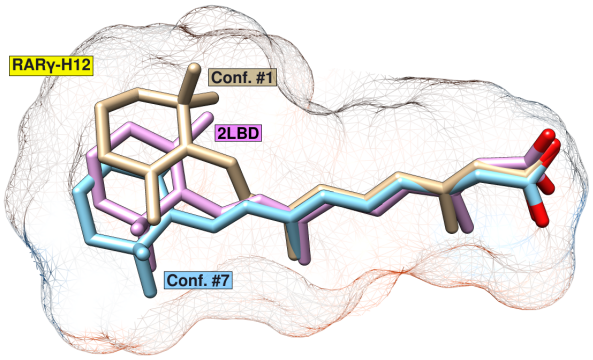

Figure 5 Structural overlay of the binding poses of ATRA in the binding pocket of RARγ (mesh) calculated using X-ray diffraction (PDB: 2LBD) and docking (conformation 1 and conformation 7). The highest scoring examples of conformations 1 and 7, respectively, show the same position of the polyene chain to that in the crystal structure.

Therefore, given the conformational flexibility of ATRA, the range and frequency of the conformations selected during docking highlighted in Figure 3, and the observation that the crystal structure of ATRA in RARγ could represent an average of a number of conformations, we can suggest that ATRA may well bind to its targets in a variety of conformations, and that at physiological temperature, each conformation may be readily accessible and interconverted while bound in the pocket.[29]

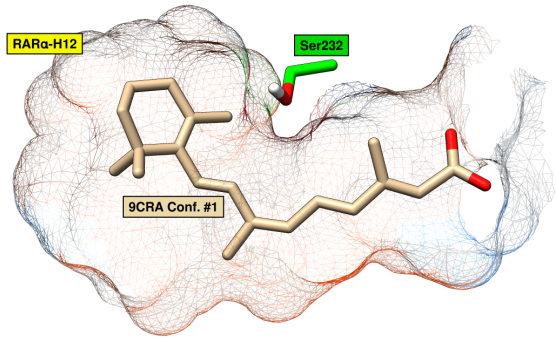

Figure 6 Highest scoring docking pose of 9CRA docked into the RARα binding pocket. The RARα-specific residue Ser232 is highlighted, demonstrating the close distance between the ring methyl and the serine side chain.

This analysis was further applied to the ATRA isomer, 9CRA, which differs from ATRA by a change in geometry along the polyene chain due to the different alkene configuration at C_9. 9CRA exhibits very similar positioning in RARα, β and γ, with a clear preference for conformation 1 in all cases. Further examination of the docking poses revealed the key role of the RARα-specific residue Ser232 in this strict discrimination against conformation 2.[40,41] As shown in Figure 6, the C_7 methyl group of the β-ionone ring of 9CRA conformation 1 must point towards Ser232 in order to accommodate the hydrophobic region in the pocket. However, due to the ease of rotation about the C_5-C_8 single bond, this methyl rotates away from a direct clash with Ser232 and towards the neighbouring Leu266 and Leu269, thus allowing the formation of strong hydrophobic contacts. While the same rotation would also occur with conformation 2, it is likely that one of the corresponding methyls of the *gem*-dimethyl moiety would still point directly at Ser232, causing an unfavourable clash that explains why the docking algorithm exclusively selected conformation 1 as most favourable in RARα.[18]

The corresponding residues in RARβ (Ala225) and RARγ (Ala234) on the other hand allow closer proximity to the single or *gem*-dimethyl groups, which explains why the docking algorithm is able to allow conformation 2 in RARβ and RARγ. Nevertheless, the much greater steric size of the *gem*-dimethyl group is likely to cause a degree of steric clashing in this tight area of the pocket, and conformation 1 is therefore still the favoured conformation.[19]

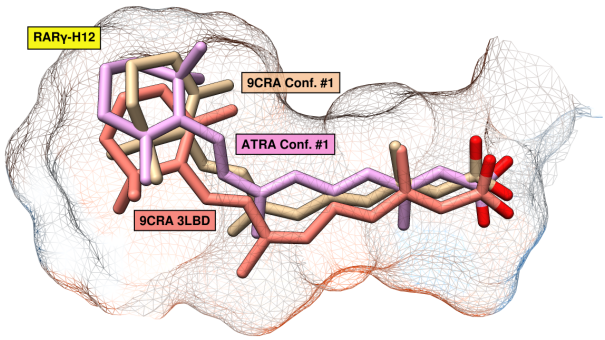

Figure 7 Highest scoring docking poses of 9CRA and ATRA overlaid with the binding pose of 9CRA in the RARγ binding pocket as determined by X-ray diffraction (PDB: 3LBD).[42] The same 9CRA conformation (9CRA conformation 1) was selected by docking as in the crystal structure, with similar positioning.

When the highest scoring docking pose of 9CRA in RARγ was compared to that of the crystal structure of 9CRA in the PDB (PDB: 3LBD), good agreement was observed (Figure 7).[42] The docked structure is positioned more towards the top of the pocket, but the two structures both adopt conformation 1 and the overall positioning is very similar, thus providing further confidence in the docking protocol (the binding pose of 9CRA in RARβ as calculated by docking was also compared to the literature crystal structure PDB: 1XDK,[15] showing equally good agreement, see ESI). When the highest scoring solution of ATRA in RARγ was superimposed, the two poses were very similar, with the exception that the β-ionone ring adopted opposite orientations, and the methyl at the carboxylate end was found to be pointed downwards, rather than upwards. Furthermore, when the positions of the carboxylate moiety between the 9CRA binding poses and that of ATRA conformation 1 were compared, clear differences were observed, dictated by their different chemical structures. Where ATRA is well placed for short, strong polar contacts between both oxygens of the carboxylate with Arg278 and Ser289, only the lower oxygen is available for hydrogen bonding to these residues in 9CRA. Similar effects were observed in RARα and β. Based on these observations, one can predict that ATRA is more suited than 9CRA for binding to the RARs, and indeed has been demonstrated experimentally (*vide infra*).[24,26] It therefore follows that based on this docking analysis, good synthetic mimics of ATRA should possess a more linear shape, which will result in the positioning of the carboxylate group closer to the polar cluster to enable effective binding interactions to the RARs.

With this knowledge in hand, we then analysed the docking solutions of ATRA-analogue EC23 with reference to the top clustering solution of ATRA in RARγ. Figure 8 shows a superposition of the highest scoring examples of ATRA conformation 1, ATRA conformation 7 and EC23, and highlights the similarity in the predicted overall positioning of these binding poses, with each exhibiting an extended, linear shape. The carboxylate of EC23 lies in an identical position to that of ATRA, allowing stabilising interactions to the polar cluster at the end of the pocket. Interestingly, the tetrahydronaphthalene hydrophobic region of EC23 essentially occupies both possible positions of the 6-membered ring of ATRA according to ATRA conformation 1 and ATRA conformation 7. By straddling both positions, interactions between the hydrophobic region of EC23 and both the upper (Leu271 & Ala234) and lower (Phe230) parts of the pocket are improved, and the two *gem*-dimethyl groups extend the reach of the ligand into all corners of the highly hydrophobic opening of the pocket.[13,42] The central acetylene linker is positioned slightly nearer the top of the pocket than ATRA, enhancing interactions with Ala234, but reducing the strength of contacts with Met272 and Ile275.[39] In comparison to ATRA, the larger benzoate polar region allows a closer positioning to the top of the pocket (Leu233), near the polar cluster, though slightly further from the bottom (Arg274 & Ile275).

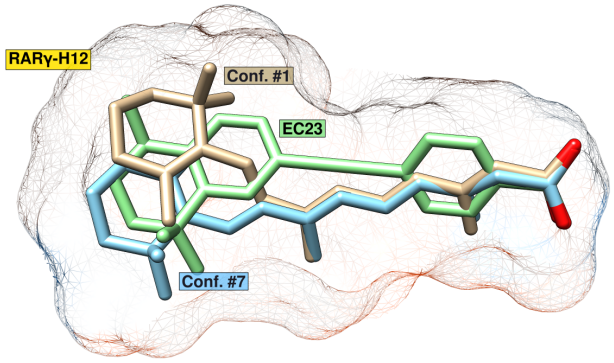

Figure 8: Structural overlay of the highest scoring docking poses of the most frequent binding conformations of ATRA, and EC23 in the binding pocket of RARγ (PDB: 2LBD) (mesh) calculated by docking simulation. The cyclohexenyl rings of ATRA conformation 1 and ATRA conformation 7 are in positions that are both straddled by the tetrahydronaphthalene hydrophobic region of EC23.

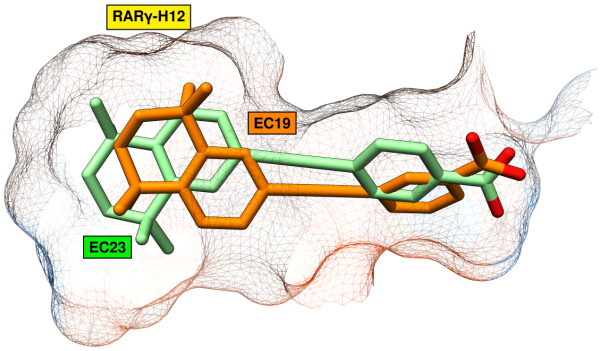

Figure 9: Structural overlay of the highest scoring docking poses of EC23 and EC19 in the binding pocket of RARγ (PDB: 2LBD) (mesh) calculated by docking. EC19 exhibits an opposite position of the hydrophobic region due to the need to accommodate the *meta*-carboxylate.

In general, according to the docking results, EC23 fills the binding pockets of RARα, β and γ more completely than ATRA, particularly in the hydrophobic areas at the opening of the pocket near H12 and next to the polar cluster at its end. The hydrophobic region of EC23 also occupies the majority of both hydrophobic region positions of ATRA conformations 1 and 7. Given that EC23 occupies the same positioning as ATRA, but interacts with more of the binding pocket, one would expect EC23 to exhibit equal, or greater, binding affinity for the RARs when compared to ATRA.

When the docking results of the *meta*-oriented carboxylate retinoid EC19 were compared to those of EC23, clear differences in binding poses were identified. The *meta*-carboxylate of EC19 was positioned toward the top of the pocket (Figure 9, in the case of RARγ). Ostensibly, this is not an unfavourable positioning, but it is one that moves one of the carboxylate oxygens further from one of the polar cluster residues (Arg278) and EC19 will likely form fewer and weaker interactions with this residue than those of EC23.[13] This alternative positioning of the carboxylate can also be seen in the binding poses in RARα and RARβ, where the increased distance (and poor angle) to the corresponding serine residue (O--H distance: 4.1, 3.9 and 3.5 Å, in RARα, β and γ, respectively) likely eliminates any hydrogen bonding interaction. Furthermore, in comparison to EC23, the highest scoring binding poses of EC19 do not form additional hydrogen bonds to the amide NH of Phe286-Ser287 in RARα and Phe279-Ser280 in RARβ, though the positioning of the carboxylate of EC19 is likely reasonable for hydrogen bonding to this amide in RARγ. These strong interactions are thought to anchor retinoids inside the binding pockets of the RARs and their absence or reduced strength in the case of EC19 may result in a reduced binding affinity.[13]

The other notable feature of the comparison between EC19 and EC23 are the differences in the positions of their hydrophobic regions. In RARγ (Figure 9) the hydrophobic regions of the two retinoids are clearly shown in two opposing orientations. While this effectively fills the same part of the pocket, the upper *gem*-dimethyl of EC19 is positioned more towards Ala234 at the top, while that of EC23 is pointed more towards to the opening of the pocket (Ile412). This EC19 positioning is likely to reduce the hydrophobic interactions with the lower left part of the pocket. In RARα, EC19 exhibits the same hydrophobic region orientation as EC23, but as with RARγ the overall positioning is biased more towards the top part of the pocket, thus likely reducing the strength of interaction with the bottom of the binding pocket.

However, as shown in Figure 10, EC23 and EC19 exhibit a much more marked difference in positioning in RARβ. In order to accommodate the *meta*-carboxylate group, the hydrophobic region of EC19 is oriented upwards towards the RARβ-specific cavity at the top of the pocket, rather than the typical retinoid positioning exemplified by EC23. This maximises hydrophobic contacts in this upper region, and EC19 is closely positioned for hydrophobic interactions with Leu262, Ile263 and Ile380, among other residues. A widely employed design strategy for the development of RARβ-specific retinoids involves the addition of sterically bulky lipophilic groups that would be positioned towards this hydrophobic cavity. Therefore, given the orientation of EC19 towards this cavity, as predicted by docking, one can predict that the molecule exhibits specificity towards RARβ.[20,43]

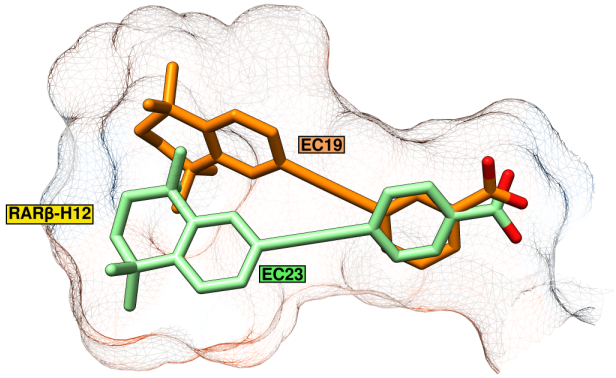

Figure 10: Structural overlay of the highest scoring docking poses of EC23 and EC19 in the binding pocket of RARβ (PDB: 1XAP) (mesh) calculated by docking simulation. The EC19 hydrophobic region is clearly oriented towards the RARβ-specific cleft at the top of the pocket

In the next step, the corresponding methyl esters of EC19 and EC23 were docked in order to confirm the key role of the carboxylate interaction with the polar cluster at the bottom of the binding pockets of the RARs.[9] In principal, with a reduced ability to form a hydrogen bond, one would expect that these esters (EC23Me, and EC19Me) exhibit reduced binding affinities. Stem-cell differentiation experiments involving EC23Me and EC19Me in our laboratories indicate significantly reduced activity *in vitro*,[44] though the level of activity may well be determined by the ease, and rate of hydrolysis to the corresponding acid and therefore, the time course of the experiment.[45]

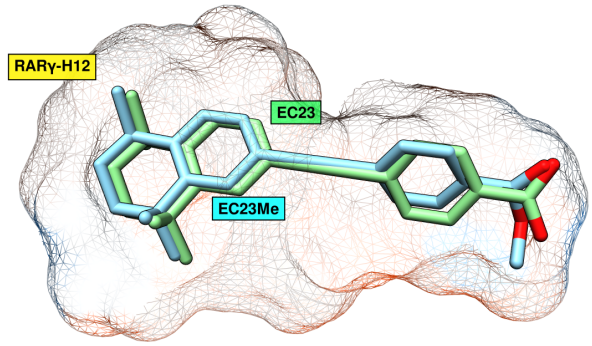

Figure 11: Comparison between the highest scoring docking poses of EC23 and EC23Me in RARγ (PDB: 2LBD) (mesh). Docking predicts the same positioning for the two ligands, but with EC23Me moved back slightly towards the entrance to the pocket in order to accommodate the methyl ester group.

Figure 11 shows a comparison of the docking poses of EC23 and EC23Me in RARγ, and highlights the effect of the ester group. In order to accommodate the ester methyl group in this tight, and relatively hydrophilic part of the pocket EC23Me has been pushed backwards towards the mouth of the pocket. While the carbonyl oxygen is still well placed for hydrogen bonding interactions with Ser289, this movement results in the ester oxygen being relatively far from Arg278 (2.7 Å) is more significant. These movements both serve to reduce the efficacy of polar contacts with the polar cluster at the end of the pocket and illicit short clashes with the polar cluster residues disrupting any opportunity for methyl ester oxygens to form hydrogen bond interactions. Similar effects are seen when comparing EC23 and EC23Me in both RARα and β, though in the case of RARβ in particular, the change in positioning is (particularly the conserved arginine).[13] Given the reduction in the number, and strength of the polar contacts in this area, one would expect that the binding affinity of EC23Me in the RARs would be reduced in comparison to EC23.

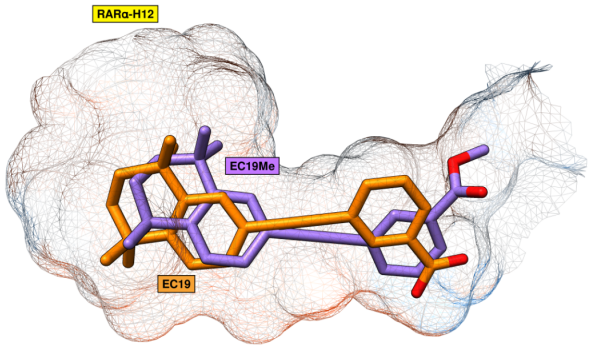

Figure 12: Comparison of the highest scoring docking poses of EC19 and EC19Me in RARα (PBD: 3KMR) (mesh). The ester of EC19Me is forced into an alternative positioning, which sacrifices hydrogen bonds with the polar cluster side chains but significantly enhances a polar contact with the main chain amide. EC19Me is positioned closer to the hydrophobic roof of the pocket.

When comparing the docking poses of EC19 and EC19Me (Figure 12) an analogous orientation of the ester away from the polar cluster was observed in RARβ and γ, but a more significant disparity in binding pose was noticed in RARα. Where EC19 lacks hydrogen bonding with the polar cluster and one of the carboxylate oxygens, EC19Me in RARα is unable to form hydrogen bonds with the Arg276 or Ser287 side-chains. However, to compensate, the carbonyl oxygen is positioned close to the Phe286-Ser287 amide link, and forms a significant (O--H distance: 1.65 Å) hydrogen bond interaction with the amide NH. As a result of this atypical positioning, the benzoate polar regions of the two exhibit opposing orientations, which in turn forces the hydrophobic region of EC19Me further towards the top part of the pocket. In comparison to EC19, the EC19Me hydrophobic region fills the RARα pocket more completely. These observations suggest that EC19Me could exhibit stronger binding to RARα than EC19 through an atypical binding pose.

The hydrogen bonding characteristics of each retinoid in RARα are compared in Figure 13 (Comparisons for RARβ/γ are shown in the ESI). Both ATRA and EC23 form electrostatic contacts involving both carboxylate oxygens and the conserved Arg and Ser side chains, and with the amide NH of the serine in RARα, β and γ. In contrast, due to the altered geometry of the ligand, 9CRA is only able to form hydrogen bonding interactions with one of its carboxylate oxygens. This situation is mirrored with EC19, where the *meta*-carboxylate is positioned such that one oxygen is far from the polar cluster. With the methyl esters, steric clashing from the methyl group and slight movement away from the polar cluster are likely to reduce the strength of hydrogen bonding interactions.

MedChemComm

ARTICLE

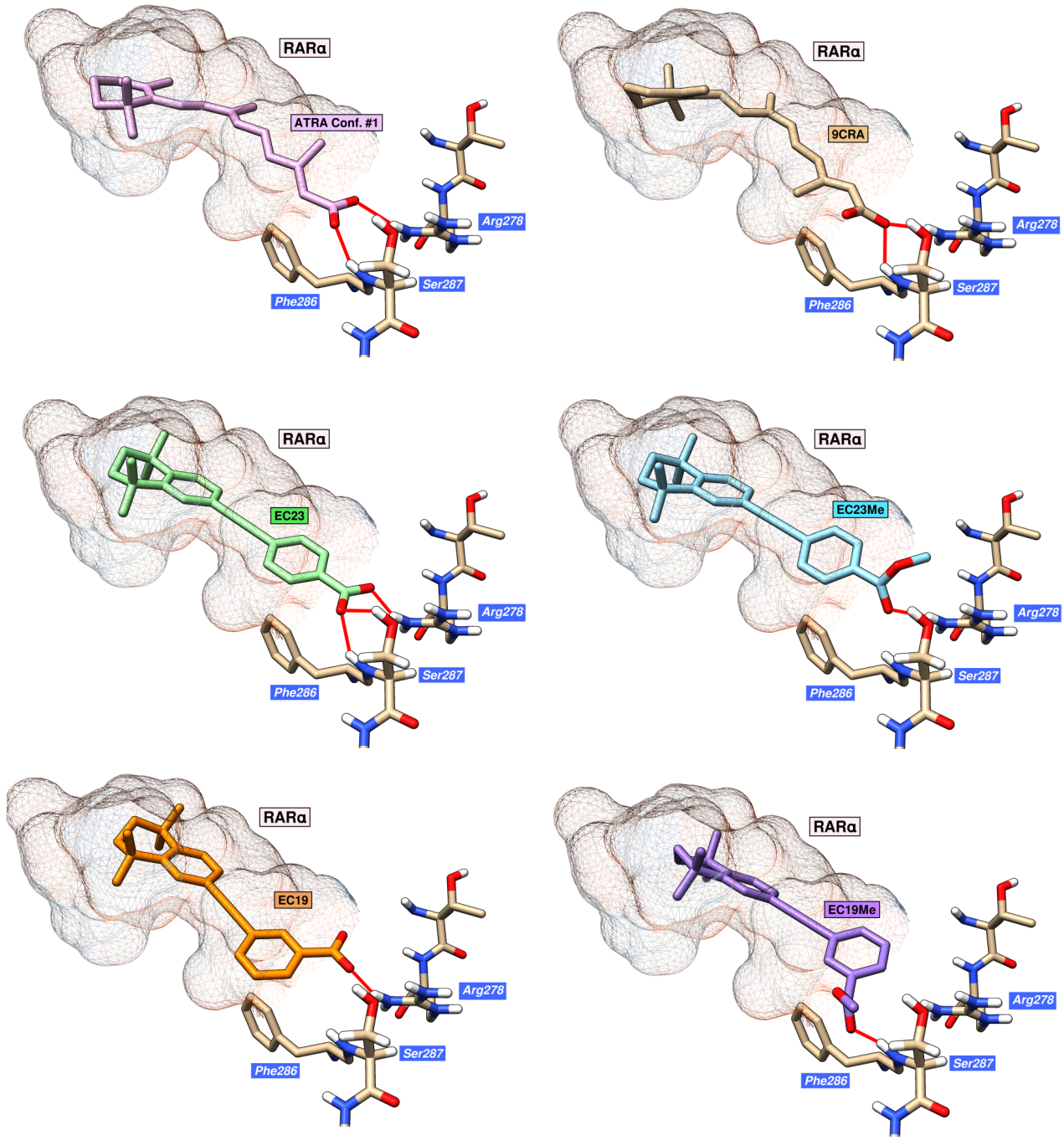

Figure 13: Hydrogen bonding characteristics of each retinoid in RARα (PDB: 3KMR) (mesh) according to the docking study. The majority of the binding pocket is represented in mesh form, the polar cluster residues are shown in a stick representation without unipolar hydrogen atoms, and the ligands are shown in a stick representation without hydrogen atoms. Both ATRA and EC23 are well placed for polar contacts via both carboxylate oxygens. Due to their positioning, only one oxygen of the carboxylates of EC19 and 9CRA are available for polar interactions. The esters EC23Me and EC19Me are positioned further from the polar cluster at the bottom pocket, and in the ester group of EC19Me is pushed towards the side of the pocket, enabling a short hydrogen bond interaction with the main chain Phe286-Ser287 amide group.

MedChemComm

ARTICLE

With the observations from docking in hand, it was possible to assess likely binding activity by analysing the overall fit of each compound to RAR LDBs with reference to structural data in the literature (Table 1).[13,20,42] This was then used to assign a predicted ranking for each retinoid that could be tested experimentally.

Table 1: Qualitative assessment of the docking calculations.[a]

Retinoid	RAR	Flexibility	Hydrophobic	Polar	Fit	Rank
ATRA	α	High	Top/Bottom	Strong	Good	2
	β	Highest	Bottom	Strong	Good	2
	γ	High	Top	Strong	Good	2
9CRA	α	Medium	Top	Medium	Medium	3
	β	Medium	Bottom	Medium	Medium	4
	γ	Medium	Top	Medium	Good	3
EC23	α	Low	Top/Bottom	Strong	Good	1
	β	Low	Bottom	Strong	Good	1
	γ	Low	Top/Bottom	Strong	Good	1
EC19	α	Low	Bottom	Medium	Poor	6
	β	Low	Top	Medium	Good	3
	γ	Low	Top	Medium	Medium	4
EC23Me	α	Low	Top/Bottom	Weak	Good	4
	β	Low	Top	Weak	Medium	5
	γ	Low	Top/Bottom	Weak	Medium	5
EC19Me	α	Low	Top/Bottom	Strong	Medium	5
	β	Low	Top	Weak	Medium	6
	γ	Low	Top/Bottom	Weak	Medium	6

[a]A ranking for each retinoid was determined based upon: 1) the observed flexibility within the binding pocket (number of binding conformations); 2) whether the retinoid interacts with the top, bottom or both areas of the pocket around H12; 3) the perceived strength of polar contacts based on the number and distance to the polar cluster; and 4) the overall fit.

The rank ordering in Table 1 suggests that EC23 is the best ligand for interacting strongly with al three RARs, followed by ATRA and 9CRA, and then some way behind is EC19, and finally the two methyl esters, EC23Me and EC19Me:

EC23 > ATRA > [$9CRA^{\alpha,\gamma}$, $EC19^{\beta}$] > [$9CRA^{\beta}$, $EC19^{\gamma}$, $EC23me^{\alpha}$] > [$EC23me^{\beta,\gamma}$, $EC19me^{\alpha}$] > [$EC19^{\alpha}$,$EC19me^{\beta,\gamma}$]

Because hydrophobic ligand-protein interactions at the LBD entrance may be important determinants of biological activity in addition to ligand-LBD affinity, we tested the rank order predictions using a binding assay which reflected these dual functions of retinoid-RAR interaction.

Receptor binding assays

The readout from the receptor binding assays was the FRET signal resulting from the association of a labelled coactivator protein with ligand-bound receptor LBD, and thus was dependent on two parameters: the affinity of ligand for the LBD and the affinity of ligand-LBD complex for the coactivator, both of which may be affected by the binding of ligand to LBD. The binding curves (Figures 14a-c) were interpreted by reference to the results of computational simulations of the effects of varying ligand/LBD and ligand-LBD/co-activator affinity (Figure 14d). This computational model showed that (1) differences in the level of the upper asymptote of the binding curves reflected differences in co-activator affinity for the ligand-LBD complex; (2) variation in affinity of ligand for the LBD which did not affect ligand-LBD coactivator affinity results in a shift in the midpoint (EC50) of the binding curve; and (3) changes in the level of the upper asymptote may result from changes in coactivator affinity for the ligand-LBD complex alone, or changes in both parameters.

According to the experimental binding curves in Figure 14a-c, all compounds except EC19 showed strong binding to RARα with EC_{50} values < 50 nM. Overall, the EC_{50} values agreed with the predicted ranking, except for the methyl esters where EC23Me was better than expected for RARβ and worse than expected for RARγ, and EC19Me was better than expected for RARβ and RARγ. EC23 had the strongest binding affinity (EC_{50} = 3.7 nM) compared to all other ligands. Upper asymptotes were highest for ATRA and EC23 compared to EC19 and 9CRA, with data for 9CRA in particular suggesting a relatively poor ability to induce co-activator recruitment. RARβ displayed a narrower range of ligand binding affinities, with EC23 having the greatest affinity (EC_{50} = 3.3 nM) and 9CRA the lowest (EC_{50} = 27.5 nM). In contrast to RARα, the upper asymptotes suggested that co-activator recruitment through RARβ was optimal for ATRA, but relatively inefficient for EC23 and 9CRA. Furthermore, EC19 had a comparably high affinity for RARβ, and ligand binding also facilitated efficient co-activator recruitment. For RARγ, ATRA and EC23 had similarly high binding affinities (EC_{50} = 14.7 nM and 16.8 nM, respectively) while EC19 the lowest (EC_{50} = 96.4 nM) and an intermediate value for 9CRA. Conversely, 9CRA had the highest asymptote suggesting a greater efficiency of co-activator recruitment through RARγ compared to the other ligands, with EC19 being the least effective. The binding affinities for the endogenous

retinoids were similar with respect to rank order to those reported elsewhere.[24,46]

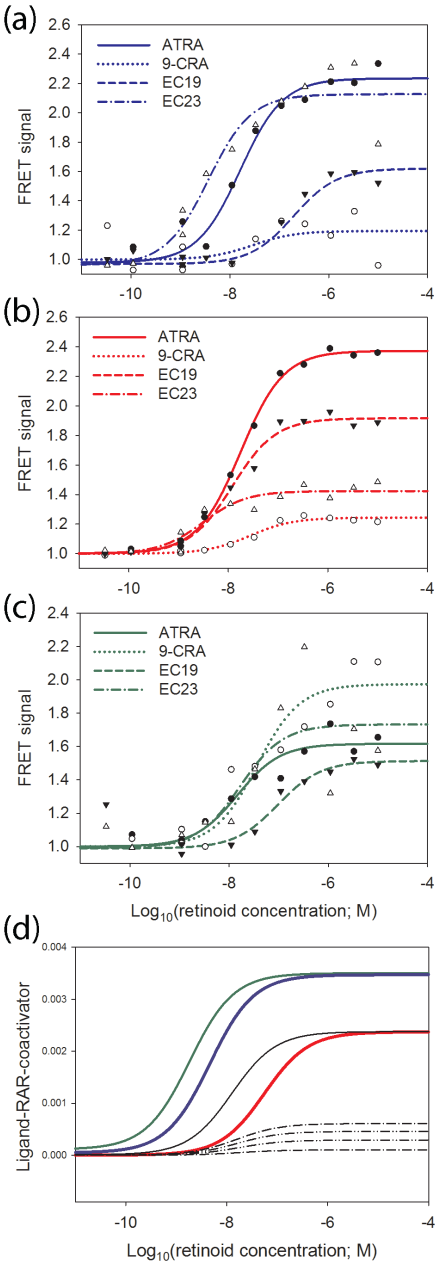

Figure 14: Three-parameter sigmoid curves fitted to binding-assay data for ATRA (●), 9-CRA (○), EC23 (△) and EC19 (▼) with RARα (a; blue lines), -β (b; red lines) and –γ (c; green lines). Ordinate axes are FRET ratio values normalised to the lower asymptote for each curve. Binding-assay simulations are shown in (d). The binding assay results show that the EC50 of EC23 for RARα was lower than ATRA, 9-CRA and EC19. For RAR-β, the EC50 of EC23 was slightly lower than ATRA, 9-CRA and EC19. With RAR-γ, the EC50 of EC23 was similar to ATRA, but 9-CRA and EC19 had a lower binding affinity. EC50 values estimated from these binding curves are given in Table 3. In the binding assay simulations with COPASI (d), the red line represents ATRA as the basal condition, the blue line models an increased affinity of co-activator for the ligand-LBD complex, the green line models an increased affinity of ligand for LBD and also co-activator for ligand-LDB, the black line represents an increase in affinity of ligand for the LBD alone, and dashed lines represent higher ligand-LBD affinity with progressively lower affinities of co-activator for ligand-LBD (see SI for details). In (d), the ordinate units are arbitrary.

Table 2: Estimated EC_{50} values (95% confidence ranges in parentheses) for ATRA, 9CRA, EC23, EC23Me, EC19 and EC19Me with RAR-α, -β and -γ.[a] n = 3).

Retinoid	RARα EC_{50}[a] (nM)	RARβ EC_{50}[a] (nM)	RARγ EC_{50}[a] (nM)
ATRA	16.0 (8.8 - 29.4)	17.6 (15.2 - 20.3)	14.7 (5.4 - 39.8)
9CRA	28.8 (0.2 - 5343.4)	27.5 (15 - 50.4)	36.7 (12.8 - 105.6)
EC23	3.7 (1.4 - 9.5)	3.3 (1.1 - 8.4)	16.8 (1.7 - 163.6)
EC19	194.0 (57.6 – 653)	12.1 (8 - 18.3)	96.4 (22.7 - 409.5)
EC23Me	20.5 (11.8 - 35.4)	14.0 (8.8 - 22.5)	847.6 (182.3 - 3940.8)
EC19Me	49.4 (7.5 - 327.2)	30.1 (18.1 – 50)	97.7 (36.1 - 264.3)

The similarity in EC_{50} values obtained for ATRA and EC23, suggest that EC23 is an excellent ATRA mimic (*vide supra*). EC23 clearly also exhibits stronger binding to RARα/β than to RARγ. The docking analysis indicated that the hydrophobic region of EC23 fills the slightly-wider RARα/β binding pockets more completely than ATRA, and this is supported by the binding assay data indicating higher binding affinity. However, while EC23 is likely a good ligand for the RARs, it may be a poor transcriptional regulator via RARβ as a result of a reduction in ability to recruit coactivator. Given that the stem cell-induction activity of EC23 is higher than that of ATRA,[9] this implies that this biological activity might be driven more through interaction of EC23 with RARα.

The lower binding affinity of 9CRA compared to ATRA would be predicted from the docking studies due to the different geometry and concomitant reduction in sites available for polar contacts, and is consistent with other reports.[24] Secondly, EC19 exhibits selectivity in binding to RARβ over RARα/γ, a result supported by the docking studies where the positioning of the hydrophobic region of EC19 provided a better fit to the RARβ-specific cleft at the top of the pocket.[47] Given this selectivity for RARβ, EC19 represents an excellent starting point for further structural optimisation. Few examples of RARβ-selective retinoids containing this *meta*-carboxylate motif exist in the literature,[3,26] and given confirmed links between RARβ and the treatment of cancers (among other conditions),[4,48–50] such a scaffold would be interesting to explore, particularly since the synthetic methodology associated with similar compounds is well understood.[9,21,51]

Finally, as might be expected, the methyl ester of EC23 exhibited reduced binding affinity compared to the parent compound, particularly for RARγ (Table 3). This is fully consistent with the docking simulation, where a movement away from the polar cluster and likely steric clashing with the methyl group would reduce electrostatic contacts. However, only EC23Me in RARγ exhibited a large decrease in binding affinity, and it may be that the hydrophobic interactions between the esters and the other hydrophobic end of the pocket are of a comparable strength to the corresponding acids in RARα/β. This was indicated in the docking studies, where only minor movements were predicted in the hydrophobic region positioning of the esters (with the exception of EC23Me in RARβ). The presence of the methyl ester in EC19Me reduced the binding affinity for RARβ (but not

RARγ), which may result from movement of the hydrophobic region and reduction in hydrogen bonding efficacy. However, as predicted from the short hydrogen bond formed between the EC19Me carbonyl oxygen and the conserved serine amide NH, EC19Me exhibited a significant increase in binding affinity for RARα compared to EC19. The esters generally exhibited reduced activity in cellular differentiation assays,[44] and the reduction in the binding affinities of these compounds compared to the corresponding acids is therefore a good rationale for this observation.

Conclusions

In this study, receptor binding assays and detailed molecular modelling have been combined to investigate the molecular basis of the biological activity of synthetic retinoids in comparison with the natural ligands. As shown previously, in cell differentiation assays, EC23 was more potent even at lower concentrations compared to the natural ligand ATRA at causing neural differentiation. In contrast, EC19 caused the development of epithelial-like structures, especially in the TERA2.cl.SP12 stem cell line.[52] These differences can be rationalised by the relative binding affinities to the three protein targets examined in this study, *i.e.* EC23 exhibited higher binding affinities for RARα, RARβ and RARγ than both ATRA and 9CRA. Conversely, EC19 had considerably lower binding affinities for both RARα and RARγ compared to EC23, but higher binding affinity for RARβ. These results may explain why EC23 can successfully induce stem cell differentiation resulting in functional neural cells in a similar manner to ATRA, but at lower concentrations. However, since EC19 bound only RARβ efficiently, this suggests that neurogenesis in response to EC23 may be driven more *via* a combination of all the RAR receptors. Interestingly, stem cell differentiation induced by EC19 gives an epithelial phenotype, and hence, this may be driven more *via* binding to RAR-β, or perhaps by less specific binding interactions at the higher concentrations required for EC19 to be effective.[53]

The molecular basis for these observations has been explored and rationalized using comprehensive molecular modelling and docking experiments. This was initially conducted with the endogenous ATRA and 9CRA to gain confidence in the docking procedure and its results. While the overall binding pose predicted for these retinoids was similar to those determined by X-ray crystallography,[13,42] for ATRA in particular we observed a wide variety of possible binding conformations. Therefore, a clustering approach was employed that examined all of the docking solutions and allowed the determination of the most frequently sampled conformations. The distribution of conformations agreed with the known structures of the respective ligand binding pockets of the three RAR isotypes, with a broader range of conformations selected in the larger RARβ compared to the more elongated binding sites in RARα/γ. Overlaying the two most frequently sampled conformations (conformation 1 and 7) with the crystal structure of ATRA bound to RARγ may indicate that the pose determined by X-ray crystallography was essentially an average of the two *in-silico* conformations, and allows us to conclude that rather than binding in one discrete conformation, ATRA likely binds in a range of conformations that lie very close to each other in energy and structure. 9CRA also exhibited some conformational flexibility, but one binding conformation was clearly favoured (conformation 1) due to the constraints of the RAR binding pocket, and indeed in RARα this was the sole binding conformation due to a steric clash between the *gem*-dimethyl group of conformation 2 with the RARα-specific residue Ser232. Interestingly, a recent study on an ATRA-receptor complex suggested that crystallography of small molecule-protein complexes may not necessarily lead to certainty about the actual small molecule conformations adopted, particularly with more conformationally flexible ligands.[29,37] An approach such as this in which all possible conformations are examined inside and outside of a protein environment substantially aids the fitting of novel ligands to crystallographic data. In turn, such an approach will further aid drug design.

These observations of the impacts of conformation effects on RAR binding were then applied to analysing the possible interactions of EC23, EC19 and their corresponding methyl esters, EC23Me and EC19Me with the RAR binding pockets. The linear structure of EC23 was found to be a good mimic for the predicted binding poses of ATRA, and further examination showed that the hydrophobic region of EC23 occupies both hydrophobic region positions of the two most frequently observed binding poses of ATRA. This 'best of both worlds' binding pose is one likely justification for the increased binding affinity exhibited by EC23 over ATRA in the *in vitro* binding assay. Furthermore, the significantly increased rigidity of EC23 is likely to be another key reason for the strong binding characteristics of this synthetic retinoid *versus* ATRA. Indeed, one would anticipate that the highly restrained structure of EC23 would experience a much-reduced loss of configurational entropy on binding when compared to ATRA/9CRA.[54] Increased binding affinity from retinoids exhibiting increased rigidity has been reported for a number of other synthetic retinoids in both the RARs and RXRs.[55,56] EC19 exhibited atypical retinoid binding according to simulation due to its *meta*-carboxylate substituent. This resulted in the sacrifice of a key hydrogen bond to the polar cluster at the bottom of the pocket, and also served to give non-ideal hydrophobic interactions around the hydrophobic region in RARα and RARγ. However, in RARβ, a predicted binding pose which significantly strengthened interactions with the RARβ-specific binding cleft was observed, and rationalized the specificity of EC19 towards RARβ in the *in vitro* assay. EC23Me and EC19Me generally exhibited reduced binding affinity compared to the corresponding acids due to the loss of hydrogen bonding interactions at the polar cluster, and also due to a movement of the hydrophobic region away from the center of the most hydrophobic part of the binding pocket. However, in RARα, docking predicted a strong hydrogen bonding contact between EC19Me and the conserved serine amide NH protein and a movement of the hydrophobic region up towards H12, which corroborated the significant increase in binding affinity for this

isotype to EC19Me in our experimental assay. This detailed level of structural analysis, summarised in Table 1, provides a relative ordering of effective binding of each ligand to the different RARs which was used to compare with receptor binding assays. The combination of molecular modelling with an experimental assay which is dependent on two separate ligand-receptor interactions strongly emphasises that biological activity is dependent on more than just ligand binding to the receptor LBD. Hence, with EC23, high LBD binding affinity does not always translate to strong coactivator recruitment and EC23 exhibited substantially greater affinity for RARβ than the other retinoids, yet produced a poor coactivator response. On the other hand, EC19 exhibited lower binding affinity but a higher coactivator response. Thus, the biological activity of natural and synthetic retinoids will depend on interactions at different levels for different receptor types, coupled with cell-type variation in abundance of different RAR types and the expression of coactivators which may determine the role of different RAR types in driving particular biological responses. Overall however, the receptor binding assays show broadly the same approximate overall rank ordering predicted by the docking studies of the different ligands with each RAR, and most importantly, the overall approach and combination of tools used in this work, in which appropriate analyses of ligand conformations in the context of *in silico* docking along with biochemical assays, predicted the effects of molecular interactions at different functional levels in a complimentary manner. Indeed, this approach is necessary to provide improved ligand design not only of new synthetic retinoids, but more widely in drug design, where a fundamental level of understanding molecular interactions, including conformational effects and their impact upon binding, is required to enable both strongly binding ligands coupled with high biological activity.

Acknowledgements

DRC thanks the EPSRC, BBSRC and High Force Research Ltd. for doctoral funding. HH thanks the Egyptian Council and Cultural Bureau for financial support.

Notes and references

1 L. J. Gudas and J. A. Wagner, *J. Cell Physiol.*, 2011, **226**, 322–330.

2 P. Chambon, *FASEB J.*, 1996, **10**, 940–954.

3 A. le Maire, S. Alvarez, P. Shankaranarayanan, A. R. de Lera, W. Bourguet and H. Gronemeyer, *Curr. Top. Med. Chem.*, 2012, **12**, 505–527.

4 L. Altucci, M. D. Leibowitz, K. M. Ogilvie, A. R. de Lera and H. Gronemeyer, *Nat. Rev. Drug Discov.*, 2007, **6**, 793–810.

5 A. Murayama, T. Suzuki and M. Matsui, *J. Nutr. Sci. Vitaminol.*, 1997, **43**, 167–176.

6 C. A. Frolik, A. B. Roberts, T. E. Tavela, P. P. Roller, D. L. Newton and M. B. Sporn, *Biochemistry*, 1979, **18**, 2092–2097.

7 J. Brtko and Z. Dvorak, *Curr. Drug Metab.*, 2011, **12**, 71–88.

8 M. A. Istrate, A. K. Nussler, M. Eichelbaum and O. Burk, *Biochem. Biophys. Res. Commun.*, 2010, **393**, 688–693.

9 V. B. Christie, J. H. Barnard, A. S. Batsanov, C. E. Bridgens, E. B. Cartmell, J. C. Collings, D. J. Maltman, C. P. F. Redfern, T. B. Marder, S. Przyborski and A. Whiting, *Org. Biomol. Chem.*, 2008, **6**, 3497–3507.

10 M. Leid, P. Kastner and P. Chambon, *Trends Biochem. Sci.*, 1992, **17**, 427–433.

11 J. Bastien and C. Rochette-Egly, *Gene*, 2004, **328**, 1–16.

12 W. Bourguet, V. Vivat, J. Wurtz, P. Chambon, H. Gronemeyer and D. Moras, *Mol. Cell*, 2000, **5**, 289–298.

13 J. P. Renaud, N. Rochel, M. Ruff and V. Vivat, *Nature*, 1995, **378**, 681–689.

14 W. Bourguet, M. Ruff, P. Chambon, H. Gronemeyer and D. Moras, *Nature*, 1995, **375**, 377–382.

15 V. Pogenberg, J.-F. Guichou, V. Vivat-Hannah, S. Kammerer, E. Pérez, P. Germain, A. R. de Lera, H. Gronemeyer, C. A. Royer and W. Bourguet, *J. Biol. Chem.*, 2005, **280**, 1625–1633.

16 T. Perlmann, K. Umesono, P. N. Rangarajan, B. M. Forman and R. M. Evans, *Mol. Endocrinol.*, 1996, **10**, 958–966.

17 J. Osz, Y. Brélivet, C. Peluso-Iltis, V. Cura, S. Eiler, M. Ruff, W. Bourguet, N. Rochel and D. Moras, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, E588-594.

18 M. Géhin, V. Vivat, J. M. Wurtz, R. Losson, P. Chambon, D. Moras and H. Gronemeyer, *Chem. Biol.*, 1999, **6**, 519–529.

19 B. P. Klaholz, A. Mitschler and D. Moras, *J. Mol. Biol.*, 2000, **302**, 155–170.

20 P. Germain, S. Kammerer, E. Pérez, C. Peluso-Iltis, D. Tortolani, F. C. Zusi, J. Starrett, P. Lapointe, J. Daris, A. Marinier, A. R. de Lera, N. Rochel and H. Gronemeyer, *EMBO Rep.*, 2004, **5**, 877–882.

21 J. H. Barnard, J. C. Collings, A. Whiting, S. A. Przyborski and T. B. Marder, *Chem. Eur. J.*, 2009, **15**, 11430–11442.

22 A. A. Levin, L. Sturzenbecker and S. Kazmer, *Nature*, 1992, **355**, 359–361.

23 R. A. Heyman, D. J. Mangelsdorf, J. A. Dyck, R. B. Stein, G. Eichele, R. M. Evans and C. Thaller, *Cell*, 1992, **68**, 397–406.

24 G. Allenby, R. Janocha, S. Kazmer, J. Speck, J. F. Grippo and A. A. Levin, *J. Biol. Chem.*, 1994, **269**, 16689–16695.

25 E. Pérez, W. Bourguet, H. Gronemeyer and A. R. de Lera, *Biochim. Biophys. Acta*, 2012, **1821**, 57–69.

26 A. R. de Lera, W. Bourguet, L. Altucci and H. Gronemeyer, *Nat. Rev. Drug Discov.*, 2007, **6**, 811–820.

27 P. F. Egea, A. Mitschler and D. Moras, *Mol. Endocrinol.*, 2002, **16**, 987–997.

28 W. Inc., .

29 X. Li, Z. Fu and K. M. Merz, *J. Comput. Chem.*, 2012, **33**, 301–310.

30 G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727–748.

31 A. le Maire, C. Teyssier, C. Erb, M. Grimaldi, S. Alvarez, A. R. de Lera, P. Balaguer, H. Gronemeyer, C. A. Royer, P. Germain and W. Bourguet, *Nat. Struct. Mol. Biol.*, 2010, **17**, 801–807.

32 C. Vasileiou, K. S. S. Lee, R. M. Crist, S. Vaezeslami, S. M. Goins, J. H. Geiger and B. Borhan, *Proteins: Struct., Funct., Bioinf.*, 2009, **76**, 281–290.

33 M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray and R. D. Taylor, *Proteins: Struct., Funct., Bioinf.*, 2003, **623**, 609–623.

34 N. J. Tatum, B. Villemagne, N. Willand, B. Deprez, J. W. Liebeschuetz, A. R. Baulard and E. Pohl, *Acta Crystallogr., Sect. C Cryst. Struct. Commun.*, 2013, **69**, 1243–1250.

35 D. K. Stafslien, K. L. Vedvik, T. De Rosier and M. S. Ozers, *Mol. Cell. Endocrinol.*, 2007, **264**, 82–89.

36 S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes and U. Kummer, *Bioinformatics*, 2006, **22**, 3067–3074.

37 Z. Fu, X. Li and K. M. Merz, *J. Chem. Theory Comput.*, 2012, **8**, 1436–1448.

38 D. M. van Aalten, B. L. de Groot, H. J. Berendsen and J. B. Findlay, *Biochem. J.*, 1996, **319**, 543–550.

39 J. Ostrowski, T. Roalsvig, L. Hammer, A. Marinier, J. E. Starrett, K. L. Yu and P. R. Reczek, *J. Biol. Chem.*, 1998, **273**, 3490–3495.

40 M. Teng, T. T. Duong, A. T. Johnson, E. S. Klein, L. Wang, B. Khalifa and R. A. S. Chandraratna, *J. Med. Chem.*, 1997, **40**, 2445–2451.

41 M. Teng, T. T. Duong, E. S. Klein, M. E. Pino and R. A. S. Chandraratna, *J. Med. Chem.*, 1996, **39**, 3035–3038.

42 B. P. Klaholz, J. P. Renaud, A. Mitschler, C. Zusi, P. Chambon, H. Gronemeyer and D. Moras, *Nat. Struct. Biol.*, 1998, **5**, 199–202.

43 P. Germain, C. Gaudon, V. Pogenberg, S. Sanglier, A. Van Dorsselaer, C. Royer, M. Lazar, W. Bourguet and H. Gronemeyer, *Chem. Biol.*, 2009, **16**, 479–489.

44 H. Haffez, Durham University, 2016.

45 S. Chen, I. M. Darling, K. L. Yu, J. E. Starrett, M. M. Mansuri, G. Whiting and K. M. Tramposch, *J. Pharm. Pharmacol.*, 1995, **47**, 626–631.

46 K.-W. Hu, X.-H. Pan, F.-H. Chen, R. Qin, L.-M. Wu, H.-G. Zhu, F.-R. Wu, J.-F. Ge, W.-X. Han, C.-L. Yin and H.-J. Li, *Int. J. Mol. Med.*, 2014, **33**, 415–422.

47 A. Johnson, E. Klein and L. Wang, *J. Med. Chem.*, 1996, **39**, 5027–5030.

48 B. Houle, C. Rochette-Egly and W. E. Bradley, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 985–9.

49 X.-C. Xu, *Cancer Lett.*, 2007, **253**, 14–24.

50 L. Altucci and H. Gronemeyer, *Nat. Rev. Cancer*, 2001, **1**, 181–193.

51 D. R. Chisholm, G.-L. Zhou, E. Pohl, R. Valentine and A. Whiting, *Beilstein J. Org. Chem.*, 2016, **12**, 1851–1862.

52 D. J. Maltman, V. B. Christie, J. C. Collings, J. H. Barnard, S. Fenyk, T. B. Marder, A. Whiting and S. A. Przyborski, *Mol. BioSyst.*, 2009, **5**, 458–471.

53 G. Clemens, K. R. Flower, P. Gardner, A. P. Henderson, J. P. Knowles, T. B. Marder, A. Whiting and S. Przyborski, *Mol. BioSyst.*, 2013, **9**, 3124–3134.

54 C. Bissantz, B. Kuhn and M. Stahl, *J. Med. Chem.*, 2010, **53**, 5061–5084.

55 M. F. Boehm, L. Zhang, B. A. Badea, S. K. White, D. E. Mais, E. Berger, C. M. Suto, M. E. Goldman and R. A. Heyman, *J. Med. Chem.*, 1994, **37**, 2930–2941.

56 J. D. Love, J. T. Gooch, S. Benko, C. Li, L. Nagy, V. K. K. Chatterjee, R. M. Evans and J. W. R. Schwabe, *J. Biol. Chem.*, 2002, **277**, 11385–11391.